

KẾT HỢP K-MEANS VỚI R^S-TREE CHO BÀI TOÁN TÌM KIẾM ẢNH THEO NGỮ NGHĨA

Lê Thị Vĩnh Thanh¹, Lê Mạnh Thanh², Văn Thế Thành³, Nguyễn Thị Uyên Nhi⁴

¹Trường Đại học Bà Rịa - Vũng Tàu

²Trường Đại học Khoa học, Đại học Huế

³Khoa Công nghệ Thông tin, Trường ĐH Sư phạm TP. HCM

⁴Trường Đại học Kinh tế, Đại học Đà Nẵng

thanhlvt@bvu.edu.vn, lmthanh@hueuni.edu.vn, thanhvt@hcmue.edu.vn, nhintu@due.udn.vn

TÓM TẮT: Trong bài báo này, kỹ thuật k-Means kết hợp với cấu trúc R^S-Tree được đề xuất để nâng cao hiệu quả tìm kiếm ảnh. Phương pháp này được thực hiện qua hai giai đoạn: (1) giai đoạn tạo cấu trúc R^S-Tree, k-Means được sử dụng để tách nút thành k-cụm; (2) giai đoạn hoàn thành cấu trúc R^S-Tree, k-Means được sử dụng tại tầng nút lá để phân cụm lại các phần tử. Bên cạnh đó, khung Ontology được xây dựng dựa vào túi từ thị giác nhằm biểu diễn ngữ nghĩa của các đối tượng và mối quan hệ giữa chúng trong ảnh. Với một ảnh đầu vào, việc tìm kiếm ảnh tương tự được thực hiện trên cấu trúc đã xây dựng. Sau đó, ảnh đầu vào được thực hiện phân lớp bằng phương pháp k-NN dựa vào tập các hình ảnh tương tự và tạo ra tập các từ thị giác. Cuối cùng, câu truy vấn SPARQL được hình thành từ tập các từ thị giác và thực hiện tìm kiếm ảnh và trích xuất ngữ nghĩa từ vựng trên cấu trúc Ontology. Trên cơ sở lý thuyết đã xây dựng, mô hình tìm kiếm ảnh theo tiếp cận ngữ nghĩa dựa trên cấu trúc phân cụm phân cấp và Ontology được đề xuất. Thử nghiệm được xây dựng trên các tập dữ liệu ảnh bao gồm: COREL, OxfordFlowers-17 và MS-COCO, kết quả thử nghiệm trong bài báo được đánh giá và thực hiện so sánh với các công trình gần đây cho thấy phương pháp đề xuất là khả thi.

Từ khóa: SBIR, Image retrieval, R^S-Tree, k-Means, Ontology.

I. GIỚI THIỆU

Việc áp dụng các kỹ thuật học máy như k-Means, k-NN, SVM, ANN và các phương pháp khác vào các cấu trúc phân cụm phân cấp để mở rộng khả năng lưu trữ và tăng tốc độ truy xuất là rất quan trọng khi xử lý cỡ dữ liệu lớn [9, 10, 16]. Hiện nay nhiều hệ thống tìm kiếm ảnh tương tự theo nội dung (CBIR) áp dụng các kỹ thuật kết hợp này để cải thiện hiệu suất về độ chính xác tìm kiếm [9, 11]. Bên cạnh đó, các hệ thống tìm kiếm hình ảnh tương tự theo tiếp cận ngữ nghĩa được phát triển để giải quyết vấn đề thu hẹp khoảng cách ngữ nghĩa giữa các đặc trưng thị giác cấp thấp của hình ảnh và ngữ nghĩa cấp cao theo suy nghĩ của người dùng. Để thực hiện điều này, nhiều phương pháp bao được áp dụng gồm các kỹ thuật học máy, mẫu ngữ nghĩa, Ontology và các phương pháp khác [19].

Tìm kiếm ảnh là một trong những thách thức trong lĩnh vực thị giác máy tính. Để cải thiện hiệu quả truy xuất và độ chính xác của quá trình tìm kiếm hình ảnh tương tự, chúng tôi đề xuất cấu trúc mới bằng việc sự kết hợp kỹ thuật k-Means với cấu trúc lưu trữ và phân cụm R^S-Tree [12]. Quá trình xây dựng cấu trúc KMRST (*k-Means R^S-Tree*) bao gồm hai giai đoạn chính. Giai đoạn đầu tiên là tạo cấu trúc R^S-Tree, trong đó thuật toán k-Means được áp dụng để tách nút thành k-cụm, giúp tối ưu hóa việc lưu trữ dữ liệu không gian đa chiều. Giai đoạn thứ hai là hoàn thành cấu trúc R^S-Tree, thuật toán k-Means tiếp tục được sử dụng tại tầng nút lá để phân cụm lại các phần tử. Sự kết hợp này cho phép tận dụng ưu điểm của cả hai phương pháp, giúp tăng cường hiệu quả về độ chính xác tìm kiếm của hệ thống truy xuất hình ảnh. Bên cạnh đó, để mô tả ngữ nghĩa của các đối tượng trong ảnh, khung Ontology được xây dựng dựa vào túi từ thị giác. Quá trình tìm kiếm tập các hình ảnh tương tự của ảnh đầu vào được thực hiện dựa trên cấu trúc lưu trữ và phân cụm dữ liệu KMRST đã xây dựng. Sau đó, ảnh đầu vào được phân lớp bằng phương pháp k-NN dựa vào tập các hình ảnh tương tự, tạo ra tập các phân lớp ngữ nghĩa. Cuối cùng, câu truy vấn SPARQL được tạo ra từ tập các phân lớp này và tìm kiếm ảnh tương tự theo tiếp cận ngữ nghĩa trên Ontology. Dựa trên các lý thuyết đã nghiên cứu, mô hình tìm kiếm ảnh theo ngữ nghĩa được đề xuất để cải thiện hiệu quả về độ chính xác tìm kiếm.

Đóng góp chính của bài báo bao gồm: (1) Kết hợp kỹ thuật k-Means với cấu trúc R^S-Tree để nâng cao hiệu quả của quá trình tìm kiếm về độ chính xác; (2) Xây dựng Ontology dựa vào túi từ thị giác để mô tả ngữ nghĩa cho các đối tượng trong ảnh; (3) Xây dựng hệ tìm kiếm ảnh theo tiếp cận ngữ nghĩa dựa trên cấu trúc KMRST và Ontology.

II. CÔNG TRÌNH LIÊN QUAN

Áp dụng các kỹ thuật học máy vào các mô hình tìm kiếm ảnh đóng góp vào việc tăng cường độ chính xác trong quá trình tìm kiếm hình ảnh. Các công trình nghiên cứu trong các thập niên gần đây đã thể hiện điều này thông qua việc kết hợp nhiều phương pháp khác nhau cho bài toán tìm kiếm ảnh, như sau:

Trong công trình [24], Younus và cộng sự (2015) đã trình bày phương pháp kết hợp mới cho việc phân cụm hình ảnh dựa trên sự kết hợp giữa thuật toán tối ưu hóa đàn hồi (*Particle Swarm Optimization* - PSO) và thuật toán phân cụm k-means. Trên cơ sở đó, phương pháp mới cho bài toán tìm kiếm ảnh dựa trên nội dung được đề xuất, phương pháp này sử dụng các màu sắc và kết cấu để biểu diễn các hình ảnh. Bốn phương pháp trích xuất đặc trưng được sử dụng bao gồm lược đồ màu, *moments* màu, ma trận đồng xuất hiện (*co-occurrence matrices*) và *wavelet moment*. Kết quả thử nghiệm đã chỉ ra rằng hệ thống đề xuất có hiệu suất vượt trội so với các hệ thống khác về độ chính xác.

Juli Rejito và cộng sự (2017) [18] đã giới thiệu phương pháp tìm kiếm ảnh dựa trên thuật toán k-Means. Phương pháp này thực hiện phân cụm các đặc trưng lược đồ histogram của ảnh. Sau quá trình phân cụm, nhóm tác giả xây dựng tập các định danh cho bộ dữ liệu hình ảnh nhằm cải thiện tốc độ tìm kiếm tập các hình ảnh kết quả [18]. Trong nghiên cứu của H.K. Maur và cộng sự (2019), các tác giả đã áp dụng kỹ thuật k-Means để phân cụm hình ảnh dựa vào các đặc trưng thị giác cấp thấp. Cùng với điều này, kỹ thuật SVM được sử dụng để phân lớp hình ảnh, từ đó xây dựng hệ thống tìm kiếm ảnh tương tự. Theo kết quả thực nghiệm, mô hình đề xuất đã nâng cao hiệu quả truy xuất hình ảnh trên các tập dữ liệu ảnh thực nghiệm [15]. Tuy nhiên, việc thực hiện phân cụm dựa trên k-Means trên các đặc trưng có thể đòi hỏi thời gian lớn để xác định các tâm cụm khởi tạo cũng như số lượng các cụm. Đồng thời, khi dữ liệu ảnh gia tăng theo thời gian, việc phân cụm và đánh chỉ mục cho hình ảnh có thể mất nhiều chi phí về thời gian và nguồn lực.

Trong một nghiên cứu gần đây [10], Annrose Joseph và các cộng sự (2021) đã trình bày phương pháp phân cụm ảnh sử dụng thuật toán tối ưu hóa đa mục tiêu k-Means moth flame optimization (KMFO). Phương pháp này nâng cao hiệu suất của thuật toán K-means bằng cách xác định số lượng cụm tối ưu và các tâm cụm bằng cách sử dụng các giá trị đa mục tiêu thu được từ phương pháp MFO. Các đặc trưng như khoảnh khắc màu sắc (*color moments*), lược đồ màu HSV, liên kết màu (*color correlogram*), ma trận đồng xuất hiện mức xám Gray Level Co-occurrence Matrix (GLCM), phép biến đổi wavelet, màu trội và các mô tả dựa trên vùng đã được sử dụng như là vectơ đặc trưng. Các thử nghiệm đã được thực hiện trên tập dữ liệu Corel 1K và cho thấy kết quả hiệu quả so với các kỹ thuật truy vấn khác.

Thêm vào đó, nhiều nhóm nghiên cứu đã áp dụng công nghệ ngữ nghĩa Ontology vào các hệ thống tìm kiếm ảnh để nâng cao hiệu suất tìm kiếm. Ví dụ, trong nghiên cứu của Manzoor U. và cộng sự (2015) [14], họ đã sử dụng hệ thống Ontology miền đề truy xuất những hình ảnh liên quan đến truy vấn của người dùng. Để thực hiện tìm kiếm, người dùng có thể nhập khái niệm dưới dạng văn bản hoặc nhập vào một hình ảnh. Liu và cộng sự (2017) [13] giới thiệu phương pháp học ngữ nghĩa dựa trên Ontology, kết nối các phân loại hình ảnh với các đối tượng trong ảnh. Trong một nghiên cứu khác [20], Sulaiman và cộng sự (2017) trình bày khung Ontology đa phương thức để mô tả các nội dung ngữ nghĩa cấp cao của hình ảnh. M. N. Asim và cộng sự (2019) [4] đã đề xuất phương pháp truy vấn thông tin dựa trên ngữ nghĩa cho dữ liệu văn bản cũng như đa phương tiện (hình ảnh, video, âm thanh). Nhóm tác giả đã sử dụng ngôn ngữ RDF để lưu trữ thông tin và truy vấn trên Ontology. Abeer Al-Mohamade và cộng sự (2020) [3] đã đề xuất những phương pháp truy vấn mới, gọi là *weight-learner*, để giải quyết vấn đề khoảng cách ngữ nghĩa giữa các đặc trưng thị giác cấp thấp của hình ảnh và ngữ nghĩa cấp cao theo suy nghĩ của người dùng.

Trên cơ sở kế thừa và khắc phục những hạn chế của các công trình liên quan, mô hình tìm kiếm ảnh kết hợp k-Means và cấu trúc R^S -Tree được giới thiệu để cải thiện hiệu quả tìm kiếm ảnh theo nội dung. Đồng thời khung Ontology được phát triển dựa trên từ thị giác để áp dụng tìm kiếm hình ảnh theo ngữ nghĩa.

III. KẾT HỢP K-MEANS VỚI CẤU TRÚC R^S -TREE

A. Xây dựng cấu trúc gom cụm

1) Cấu trúc R^S -Tree

Cấu trúc của R^S -Tree [12] bao gồm một nút gốc, một tập hợp các nút trong và một tập hợp các nút lá. Mỗi nút trên cấu trúc R^S -Tree là một khối cầu. Trong quá trình xây dựng cây, dữ liệu đa chiều được phân chia thành các nhóm khối cầu dựa trên các đặc trưng của chúng. Các khối cầu nút trong chứa các thông tin liên kết đến các khối cầu nhóm con, trong khi các khối cầu nút lá chứa các nhóm dữ liệu vectơ đặc trưng ảnh. Khi tìm kiếm ảnh tương tự, quá trình duyệt cây bắt đầu từ khối cầu nút gốc và tiếp tục xuống tới các khối cầu nút con phù hợp dựa vào độ đo tương tự, cho đến khi đạt đến nút lá gần nhất với ảnh cần tìm kiếm. Tổ chức cấu trúc R^S -Tree thành cây phân cụm dữ liệu đa chiều giúp đạt hiệu quả trong việc tìm kiếm ảnh tương tự, đồng thời đảm bảo tính cân bằng của cây để tối ưu hóa quá trình tìm kiếm và truy xuất dữ liệu.

R^S -Tree là cấu trúc cây được phát triển dựa trên cấu trúc SS-Tree [23]. Các phương pháp tính tâm cho cả nút lá và nút trong của khối cầu trên R^S -Tree tương tự với phương pháp của SS-Tree. Tuy nhiên, R^S -Tree đã được cải tiến với một số điểm mới như sau: (1) Tạo khối cầu không gian để lưu trữ các vectơ đặc trưng của hình ảnh, giúp tăng cường khả năng lưu trữ và quản lý dữ liệu; (2) Cải tiến thuật toán tách nút dựa trên độ lệch sai biệt giữa các phân tử, giúp tăng cường tính chính xác của việc quá trình gom cụm dữ liệu; (3) Đưa ra ngưỡng θ để làm bộ lọc và phân chia các vectơ dãn trung vào các nhóm dữ liệu tương tự; (4) Kết hợp cả tìm kiếm láng giềng gần nhất và tìm kiếm vùng không gian, tận dụng hai phương pháp này để thực hiện tìm kiếm ảnh tương tự một cách hiệu quả.

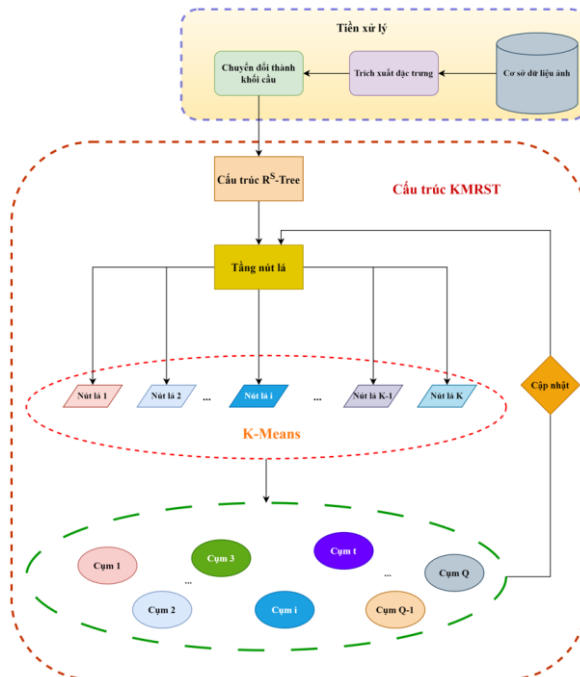
Những cải tiến này giúp R^S -Tree trở thành phương pháp hiệu quả trong bài toán tìm kiếm ảnh tương tự, giúp cải thiện độ chính xác tìm kiếm ảnh.

2) Đề xuất cấu trúc KMRST

Trong bài báo này, cấu trúc KMRST được xây dựng và phát triển dựa trên R^S -Tree. Cấu trúc KMRST đã được cải tiến nhằm giải quyết vấn đề về việc tách nút xảy ra thường xuyên, điều này ảnh hưởng đến hiệu suất tìm kiếm chưa cao. Các cải tiến trong cấu trúc KMRST bao gồm:

Đầu tiên, khi một nút lá tràn, có nghĩa nút lá này có số phần tử tối đa cho phép, thuật toán tách nút được thực hiện. Thay vì nút tràn được tách cố định thành hai nút con, phương pháp k-Means để thực hiện tách nút tràn thành k cụm ($k \geq 2$). Mục tiêu của phương pháp này là tối ưu hóa độ tương tự của các đối tượng không gian trong một cụm và đồng thời giảm thiểu độ tương tự và hạn chế chồng lấp không gian giữa các cụm khác nhau. Các phần tử trong nút tràn sẽ được chia thành k cụm dựa trên độ đo tương đồng.

Thứ hai, sau khi xây dựng cấu trúc R^S -Tree, thuật toán k-Means được sử dụng tại tầng nút lá của cây. Mỗi khối cầu của tầng nút lá có một phần tử tâm đại diện cho nút lá tương ứng. Các điểm tâm này được sử dụng làm các tâm khởi tạo ban đầu cho thuật toán k-Means. Số lượng điểm tâm ban đầu bằng với số lượng nút lá đã được tạo ra. Sau đó, các điểm dữ liệu tại tầng nút lá được phân bố lại theo thuật toán k-Means. Sau khi hoàn tất quá trình phân bố, các nút lá mới được cập nhật lại vào cấu trúc cây. Đối với những nút lá có số lượng điểm dữ liệu vượt quá ngưỡng cho phép, thuật toán k-Means được thực hiện để chia tách nút lá đó thành d cụm, với d là số nguyên nhỏ nhất lớn hơn giá trị T . Ở đây, T là tỷ lệ số phần tử dữ liệu trong lá so với số phần tử dữ liệu tối đa mà một nút lá có thể chứa. Cấu trúc KMRST được minh họa như Hình 1.



Hình 1. Cấu trúc KMRST kết hợp k-Means và R^S -Tree

Việc kết hợp phát triển cấu trúc lưu trữ dữ liệu với các phương pháp học máy trong lĩnh vực tìm kiếm ảnh có thể cải thiện hiệu suất tìm kiếm, trích xuất thông tin chính xác hơn và mở rộng khả năng tìm kiếm dựa trên ngữ nghĩa.

B. Các thuật toán xây dựng cấu trúc KMRST

1) Thuật toán tách nút sử dụng thuật toán k-Means

KMRST là cấu trúc cây cân bằng động phát triển theo chiều từ lá dưới lên gốc. Việc nút N_i bị tràn được xử lý bằng cách phân chia $M + 1$ phần tử vào k nút ($k \geq 2$). Thuật toán k-Means được sử dụng để thực việc phân hoạch $M + 1$ phần tử thành k nhóm. Sau đó mỗi nhóm sẽ được gán vào mỗi nút lá trên cây KMRST có cùng cha với nút lá được tách. Việc lựa chọn thuật toán k-Means vì ba lý do sau: (1) Hiệu quả về thời gian thực thi; (2) Độ chính xác cao; (3) Không phụ thuộc vào thứ tự của các đối tượng. Thuật toán tách nút lá được trình bày như sau:

Thuật toán 1. Tách nút sử dụng thuật toán k-Means

Input: Nút lá cần tách S_L

Output: k nút sau khi được tách ($k \geq 2$)

Function: **SplitLeafKMRST**(S_L)

- 1 **Begin**
- 2 Phân chia tập $M + 1$ phần tử thành k nút sử dụng phương pháp K-Means;
- 3 Khởi tạo số cụm với $k = 2$;
- 4 Áp dụng k-Means trên tập $M+1$ phần tử để phân chia thành k cụm;
- 5 Tính toán vùng diện tích chồng lấp giữa k vùng không gian;
- 6 $Max = OverlapAera(k)$;
- 7 $k_{opt} = k$;
- 8 For $k = 3$ to k_{max}

```

9      Áp dụng k-Means trên tập M+1 phần tử để phân chia thành k cụm;
10     Tính toán  $OverlapAera(k)$ ;
11     If  $OverlapAera(k) < Max$  then
12          $Max = OverlapAera(k)$ ;
13          $k_{opt} = k$ ;
14     For  $k = 1$  to  $k_{opt}$ 
14     Gán các phần tử của k cụm vào k nút.
15     End

```

2) Thuật toán tạo cấu trúc Improved dựa trên sự kết hợp k-Means với R^S-Tree

Để nâng cao khả năng lưu trữ và gom cụm dữ liệu đáp ứng cho các hệ thống tìm kiếm ảnh. Thuật toán xây dựng cấu trúc KMRST được hình thành bởi sự kết hợp thuật toán k-Means tại tầng nút lá của cấu trúc R^S-Tree gồm các bước được mô tả như sau:

Bước 1. Lần lượt đưa các vectơ đặc trưng ảnh vào cấu trúc R^S-Tree theo các nguyên tắc tạo cấu trúc R^S-Tree.

Bước 2. Áp dụng phân cụm k-Means cho tập hình ảnh được thu được từ tầng nút lá của cấu trúc R^S-Tree.

Bước 3. Số lượng cụm cần tạo bằng chính số nút lá được tạo ra trên cấu trúc R^S-Tree ở bước 1 (đặt là K).

Bước 4. K tâm cụm khởi tạo chính là K tâm của K khối cầu nút lá.

Bước 5. Tập dữ liệu được phối phối theo thuật toán k-Means dựa vào K tâm khởi tạo.

Bước 6. Mỗi phần tử dữ liệu xác định điểm tâm gần nhất của nó.

Bước 7. Sau mỗi lần hoàn thành phân phối tâm mới được cập nhật.

Bước 8. Lặp lại từ Bước 5 đến Bước 7 cho đến khi thu được K-cụm với tâm cố định.

Bước 9. Duyệt các cụm có số số lượng phần tử lớn hơn ngưỡng cho phép. Thực hiện k-Means một lần nữa cho các nút này với số cụm bằng d.

Thuật toán 2. Tạo cấu trúc KMRST

Input: Tập dữ liệu ảnh

Output: cấu trúc KMRST

Begin

Tạo cấu trúc R^S-Tree theo nguyên tắc tạo cây như trong [];

$K = count(\text{số nút lá được tạo ra từ cấu trúc R}^S\text{-Tree})$;

Áp dụng k-Means trên tập dữ liệu tại tầng nút lá để phân chia thành K cụm;

For $i = 1$ to K **do**

If($count(\text{cụm } i) > M$) **then**

Áp dụng k-Means cho cụm i với $k = d$;

Endif

Endfor

Gọi P là tổng số cụm được tạo ra sau quá trình k-Means

For $i = 1$ to P **do**

Gán P cụm này vào P nút cha tương ứng

Endfor

Return KMRST;

End

IV. TÌM KIẾM ẢNH KẾT HỢP KMRST VÀ ONTOLOGY

A. Tạo Ontology dựa vào túi từ thị giác

1) Mô hình tạo Ontology

Dữ liệu hình ảnh không ngừng gia tăng, vì vậy việc xây dựng và phát triển Ontology cho tập dữ liệu hình ảnh là một nhiệm vụ đầy thách thức. Miền dữ liệu thường xuyên thay đổi trong các bộ dữ liệu ảnh lớn cũng như các mối quan hệ phức tạp giữa các đối tượng trong ảnh nên đòi hỏi thời gian và nguồn nhân lực lớn để thực hiện. Để làm đa dạng hoá ngữ nghĩa hình ảnh, tiến trình xây dựng khung Ontology phải thoả mãn yêu cầu bổ sung được nguồn dữ liệu mới vào trong cấu trúc Ontology.

Quy trình phát triển Ontology cho các bộ ảnh đa đối tượng và đơn đối tượng được đề xuất như trong Hình 2, sử dụng tập ảnh COREL, OxfordFlowers17 (OF-17), MS-COCO. Các đối tượng trong các hình ảnh được phân lớp và đưa vào các túi từ. Sau đó, các phân lớp, các cá thể đối tượng trong hình ảnh, các cá thể hình ảnh và thuộc tính của chúng được phân phối vào ontotoly nhằm mục đích nâng cao hiệu quả tìm kiếm ảnh.

3) Thuật toán phân phối túi từ thị giác vào Ontology

Bằng cách sử dụng túi từ thị giác, ta có thể rút trích và phân loại các đối tượng hình ảnh, sau đó gắn chúng vào Ontology theo các nhóm tương ứng. Các ảnh gốc đi kèm với ID ảnh và thuộc tính của chúng, các mô tả ảnh được gắn cho từng đối tượng ảnh tương ứng trong Ontology. Cách thức đưa các đối tượng và hình ảnh vào Ontology được mô tả như sau:

Thuật toán 4. Phân bố hình ảnh vào Ontology

ĐẦU VÀO: Túi từ B, Ontology *OB*)

ĐẦU RA: Ontology *OB* sau khi được phân bố.

Begin

for (Duyệt lần lượt từng túi từ) **do**

for (duyệt các thành phần trong túi từ) **do**

 Tạo bộ ba (Subject, Predicate, Object);

 Đưa bộ ba vào Ontology *OB*;

Endfor

Endfor

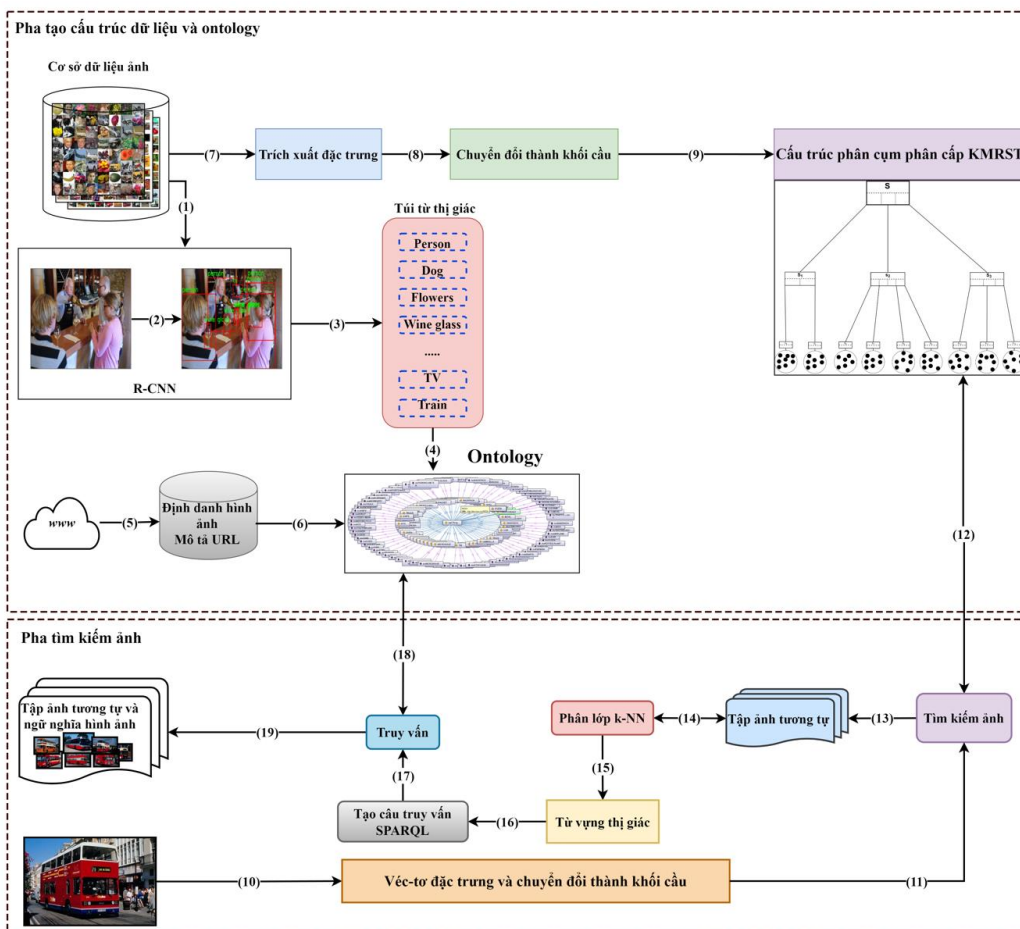
Return *OB*;

End.

Trong Thuật toán 4, Ontology được làm phong phú thêm trên hai khía cạnh bao gồm cả các đối tượng trong ảnh và các ảnh ban đầu. Đối với mỗi túi hình ảnh là một tập hợp các hình ảnh đối tượng, mỗi hình ảnh có một ID đối tượng và kết nối đến một hình ảnh gốc. Vì vậy, khi làm phong phú Ontology, các ảnh đặc trưng được thêm vào theo lớp biểu diễn và các ảnh gốc tương ứng với các đối tượng ảnh này cũng được thêm vào Ontology. Như vậy, Ontology sẽ lưu trữ cụm túi từ bao gồm các ảnh đặc trưng được liên kết với ảnh gốc để có thể lấy ra một tập ảnh tương tự từ ảnh tìm kiếm đầu vào.

B. Kiến trúc quy trình tìm kiếm ảnh

Trong phần này, kiến trúc hệ thống tìm kiếm ảnh theo tiếp cận ngữ nghĩa được giới thiệu dựa trên sự kết hợp kỹ thuật k-Means với cấu trúc R^S-Tree và Ontology được mô tả như trong Hình 4. Trong mô hình này, gồm 2 pha: pha tạo cấu trúc lưu trữ dữ liệu và pha tìm kiếm ảnh.



Hình 4. Mô hình tìm kiếm ảnh theo ngữ nghĩa KMRSTonIR

Các bước thực hiện mô hình được trình bày chi tiết như sau:

Pha tiền tạo cấu trúc lưu trữ dữ liệu gồm 2 giai đoạn:

Giai đoạn 1: tạo Ontology

Bước 1. Các hình ảnh trong cơ sở dữ liệu được phát hiện đối tượng, thực hiện trích xuất đặc trưng và phân lớp đối tượng bằng mạng nơ-ron học sâu (1, 2).

Bước 2. Tạo tuple từ thị giác của từng vùng đối tượng của mỗi hình ảnh, mỗi tuple từ thị giác mô tả một cụm các hình ảnh tương tự và phân bố từng nhóm hình ảnh này vào trong một đối tượng đại diện của Ontology (3).

Bước 3. Từ các tuple từ thị giác, Ontology xây dựng để lưu trữ cho từng cụm hình ảnh trên từng tuple từ; đồng thời các mô tả ngữ nghĩa, thuộc tính của các đối tượng và hình ảnh được đưa vào trong Ontology (4, 5, 6).

Giai đoạn 2: Tạo cấu trúc lưu trữ dữ liệu KMRST

Bước 1. Trích xuất đặc trưng cấp thấp của các ảnh trong cơ sở dữ liệu ảnh.

Bước 2. Chuyển đổi các đặc trưng ảnh về dạng khối cầu.

Bước 3. Lần lượt đưa các khối dữ liệu vào trong cấu trúc KMRST theo các quy tắc đã đề xuất.

Pha tìm kiếm ảnh tương tự: gồm hai giai đoạn

Giai đoạn 1: Tìm kiếm ảnh theo nội dung

Bước 1. Từ ảnh đầu vào, hệ thống thực hiện trích xuất và chuyển đổi về dạng khối cầu (10).

Bước 2. Thực hiện tìm kiếm trên cấu trúc KMRST (11, 12).

Bước 3. Trả về tập ảnh tương tự theo nội dung với ảnh tìm kiếm (13).

Giai đoạn 2: Tìm kiếm ảnh theo ngữ nghĩa

Bước 1. Từ kết quả tập ảnh tương tự được tạo ra ở giai đoạn 1, phương pháp k-NN được sử dụng để phân lớp ảnh đầu vào tạo ra tập từ vựng thị giác (14, 15).

Bước 2. Câu lệnh SPARQL được tạo ra dựa trên tập các phân lớp của ảnh đầu vào và được thực thi trên Ontology đã xây dựng (16, 17, 18).

Bước 3. Kết quả trả về *topk* hình ảnh tương tự nhất với ảnh tìm kiếm, đồng thời các khái niệm ngữ nghĩa được mô tả (19).

Mô hình tìm kiếm ảnh tương tự KMRSTonIR kết hợp tìm kiếm theo nội dung và theo ngữ nghĩa dựa trên sự kết hợp cấu trúc gom cụm KMRST và Ontology nhằm nâng cao độ chính xác tìm kiếm. Việc sử dụng Ontology để lưu trữ và mô tả ngữ nghĩa dữ liệu hình ảnh trong bài toán tìm kiếm ảnh tương tự mang lại nhiều ý nghĩa quan trọng: (1) Ontology giúp xác định và định nghĩa các khái niệm trong miền dữ liệu hình ảnh cụ thể; (2) Ontology giúp tăng hiệu năng của các hệ thống tìm kiếm ảnh tương tự theo tiếp cận ngữ nghĩa; (3) Ontology cho phép mô tả các mối quan hệ phức tạp giữa các đối tượng trong dữ liệu hình ảnh, điều này giúp hỗ trợ truy vấn phức tạp và chi tiết hơn từ phía người dùng; (4) Ontology giúp tối ưu hóa quy trình tìm kiếm ảnh tương tự bằng cách cung cấp một khung có cấu trúc và tổ chức về dữ liệu hình ảnh; (5) Ontology cung cấp cơ chế chung để tích hợp dữ liệu từ nhiều nguồn khác nhau, bao gồm cơ sở dữ liệu hình ảnh, hệ thống thông tin và tri thức chuyên môn.

Phương pháp sử dụng Ontology để lưu trữ và mô tả ngữ nghĩa dữ liệu hình ảnh trong bài toán tìm kiếm ảnh tương tự giúp cải thiện tính chính xác, hiệu suất và linh hoạt của hệ thống truy vấn ảnh. Thực nghiệm mô hình này được trình bày chi tiết trong Mục V.

C. Các thuật toán tìm kiếm ảnh

1) Thuật toán tìm kiếm ảnh theo nội dung

Quá trình tìm kiếm ảnh theo mô hình đề xuất là sự kết hợp từ nội dung đến ngữ nghĩa. Quá trình tìm kiếm ảnh theo nội dung được thực hiện trên cấu trúc phân cụm dữ liệu KMRST đã xây dựng. Khi tìm kiếm, hệ thống duyệt từ nút gốc, và tiếp tục đi theo hướng có độ đo tương đồng nhất về khoảng cách Euclid và vùng giao nhau nhiều nhất với ảnh tìm kiếm cho đến khi tìm thấy nút lá chứa các ảnh tương tự. Thuật toán tìm kiếm ảnh dựa trên các nội dung đặc trưng cấp thấp được mô tả như sau:

Thuật toán 5. Tìm kiếm ảnh theo nội dung trên cấu trúc KMRST

Input: Ảnh truy vấn I_Q , cấu trúc KMRST.

Output: tập ảnh tương tự và tập các phân lớp của ảnh tìm kiếm.

Function: **ImprovedRSTIR**($Root, I_Q$)

```

1   begin
2       if ( $Root == null$ ) then
3           trả về giá trị  $null$ ;
4       else
5           if ( $Root$  khác với nút lá) then
6               Tính toán độ tương đồng dựa trên độ đo khoảng cách Euclid và chọn nhánh phù hợp;
```

```

7       Giả sử  $S_N$  là nhánh được chọn;
8       ImprovedRSTIR ( $S_N, I_Q$ );
9       else
10      Trả về tập ảnh tương tự  $SI$  của ảnh  $I_Q$ ;
11      Phân lớp ảnh sử dụng phương pháp  $k$ -NN cho tập ảnh tương tự  $SI$ ;
12       $C = kNNclassify(SI, k)$ ;
13      Return  $\langle SI, C \rangle$ ;
14  End

```

2) Thuật toán tìm kiếm ảnh theo ngữ nghĩa

Quá trình tìm kiếm ảnh trên Ontology được thực hiện theo các bước sau:

Bước 1. Phân lớp ảnh đầu vào

Mỗi ảnh đầu vào trải qua quá trình tìm kiếm trên cấu trúc KMRST, sau đó trả lại tập hình ảnh có nội dung tương tự. Tiếp theo, kỹ thuật phân lớp k-NN được áp dụng để phân loại ảnh đầu vào dựa trên tập hình ảnh tương tự đã thu thập. Kết quả của quá trình phân loại này tạo ra bộ từ vựng thị giác. Sử dụng bộ từ vựng thị giác này, chúng ta tạo ra các câu truy vấn SPARQL để thực hiện truy vấn trên Ontology và tìm kiếm các hình ảnh tương tự.

Bước 2. Truy vấn ảnh tương tự trên Ontology

Với mỗi ảnh đầu vào, tập các phân lớp được xác định dựa trên nhãn ảnh đối tượng. Từ đó, câu lệnh SPARQL được khởi tạo ở Bước 1 và sau đó thực thi câu lệnh truy vấn trên Ontology đã xây dựng, cuối cùng trả về tập kết quả ảnh tương tự theo ngữ nghĩa tương ứng. Thuật toán tìm kiếm ảnh trên Ontology được mô tả như sau:

Thuật toán 6: Thuật toán tìm kiếm và trích xuất ngữ nghĩa hình ảnh trên Ontology.

Đầu vào: Ảnh truy vấn I_q

Đầu ra: Tập ảnh tương tự SSI ;

Begin

Khởi tạo tập nhãn lớp: $Label = \emptyset$;

Phân lớp ảnh tìm kiếm I_q sử dụng kỹ thuật phân lớp k-NN để có tập nhãn $Label$;

for (Duyệt các nhãn lớp thuộc tập $Label$) **do**

Câu lệnh SPARQL được tạo;

Endfor

Thực thi câu truy vấn trên Ontology OB trả về tập ảnh tương tự SSI ;

Return SSI ;

End.

Trên cơ sở các thuật toán đã có, thuật toán truy vấn ảnh được thực hiện dựa trên quá trình truy xuất nhãn lớp và tạo câu truy vấn SPARQL. Với đầu vào là hình ảnh tìm kiếm và Ontology, thuật toán truy vấn trả về tập kết quả gồm các hình ảnh tương tự với ảnh đầu vào.

Đầu tiên, ảnh đầu vào được phân lớp bằng phương pháp k -NN dựa vào tập ảnh tương tự theo nội dung được truy vấn trên cấu trúc KMRST. Từ tập nhãn lớp này, câu truy vấn SPARQL được tạo ra dựa trên ngôn ngữ bộ ba. Câu truy vấn SPARQL này thực thi tìm kiếm trên Ontology theo ngôn ngữ bộ ba RDF để từ đó trích xuất các hình ảnh thỏa điều kiện của bộ ba này. Vì vậy tập các dữ liệu meta-data trên Ontology được truy vấn làm cơ sở cho quá trình tra cứu hình ảnh gốc ban đầu tương tự với ảnh đầu vào.

V. THỰC NGHIỆM

A. Môi trường và ứng dụng thực nghiệm

Hệ truy vấn KMRSTOnIR được xây dựng dựa trên nền tảng dotNET Framework 4.8, ngôn ngữ lập trình C#. Quá trình tạo cấu trúc dữ liệu và xây dựng Ontology được thực hiện trên máy PC CPU 2.3 GHz 8-core 9th-generation Intel Core i9, 16 GB 2666 MHz memory, 1TB flash storage. Quá trình tìm kiếm ảnh được thực nghiệm trên máy PC CPU Intel Core i7-6500U CPU @ 2.50 GHz, 8.0 GB RAM, hệ điều hành Windows 10 Pro 64 bit. Tập dữ liệu được sử dụng trong thực nghiệm là bộ dữ liệu ảnh COREL, OxfordFlowers17, MS-COCO với 118.287. Thông tin chi tiết của các bộ dữ liệu thực nghiệm được trình bày như Bảng 1.

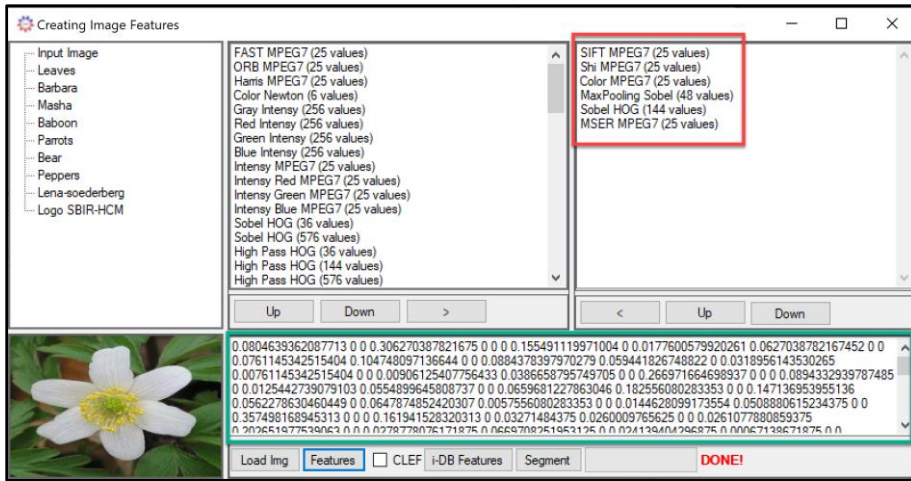
Bảng 1. Các bộ ảnh được sử dụng để thực nghiệm cho mô hình đề xuất

STT	Tên tập ảnh	Số lượng ảnh	Số lượng lớp ảnh	Kích thước
1	COREL	1.000	10	30,3 MB
2	Oxford-Flowers 17	1.360	17	60.5 MB
3	MS-COCO	118.287	80	17,9 GB

1. Ứng dụng trích xuất đặc trưng ảnh

Trong bài báo này, mỗi đối tượng ảnh được trích xuất vectơ đặc trưng tổ hợp từ các đặc trưng khác nhau gồm: Đặc trưng vị trí (*location*), đặc trưng màu sắc (*color*), đặc trưng bề mặt, vân ảnh (*texture*), đặc trưng hình dạng (*shape*).

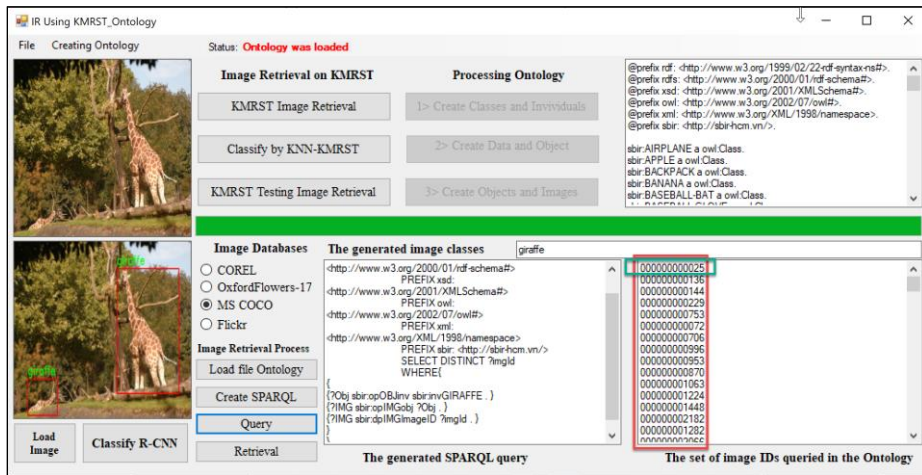
Các phương pháp trích xuất đặc trưng cấp thấp được kết hợp gồm: đặc trưng màu sắc MPEG7 (25 giá trị); đặc trưng vị trí Shi-tomasi MPEG7 (25 giá trị); đặc trưng vị trí SIFT MPEG7 (25 giá trị); đặc trưng vị trí MSER MPEG7 (25 giá trị); đặc trưng vị trí và kết cấu MaxPooling Sobel (48 giá trị); đặc trưng hình dạng và kết cấu Sobel (144 giá trị). Số đặc trưng của hình ảnh được trích xuất là 292 chiều. Việc kết hợp các phương pháp trích xuất đặc trưng gồm Shi-tomashi, SIFT, SURF (MSER) giúp nâng cao độ chính xác [5]. Ứng dụng trích xuất đặc trưng được thể hiện trong Hình 5.



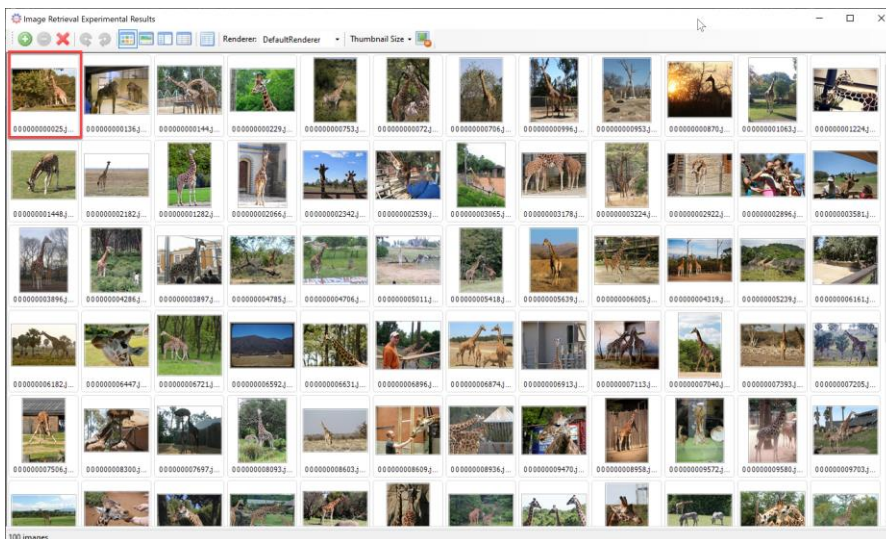
Hình 5. Trích xuất đặc trưng cấp thấp của hình ảnh

2. Ứng dụng tìm kiếm ảnh tương tự

Từ các thành phần của mô hình tại Mục III, các kết quả ứng dụng thực nghiệm được xây dựng trên ba bộ ảnh COREL, Oxford Flower 102 và MS-COCO. Các kết quả thực nghiệm được trình bày trong các Hình 7-8.



Hình 6. Tìm kiếm ảnh trên hệ RCNNOnIR của bộ ảnh MS-COCO



Hình 7. Kết quả tìm kiếm của ảnh “00000000025”.jpg của bộ ảnh MS-COCO

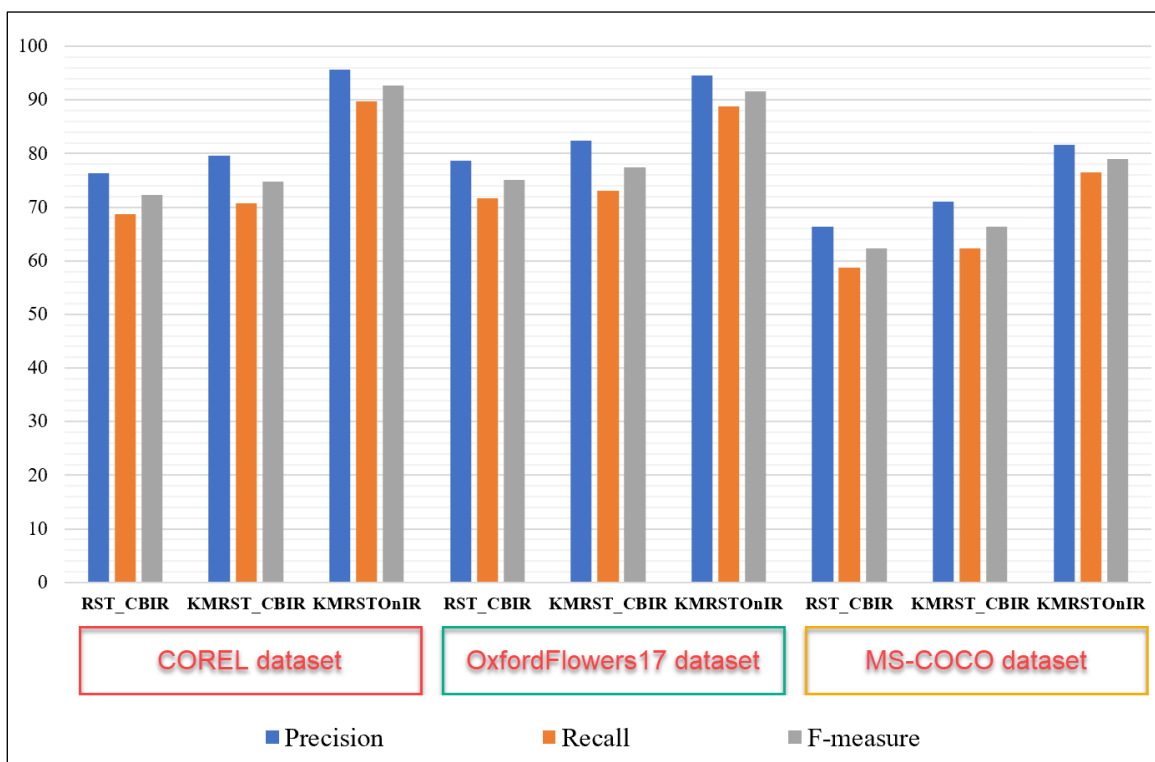
B. Phân tích và đánh giá thực nghiệm

Các giá trị hiệu suất đánh giá quá trình tìm kiếm ảnh bao gồm các giá trị sau: Precision (độ chính xác), Recall (độ phủ), F-Measure (độ đo dung hòa) và Query time (thời gian tìm kiếm). Kết quả tìm kiếm ảnh giữa các phương pháp RST_CBIR (tìm kiếm ảnh theo nội dung trên cấu trúc R^S-Tree), KMRST_CBIR (tìm kiếm ảnh theo nội dung trên cấu trúc kết hợp k-Means với R^S-Tree), KMRSTOnIR (tìm kiếm ảnh theo ngữ nghĩa theo mô hình đề xuất) được mô tả trong Bảng 2. Đồng thời, biểu đồ so sánh hiệu suất của quá trình tìm kiếm ảnh trên các mô hình của các tập dữ liệu ảnh COREL, OxfordFlowers-17 và MS-COCO được trình bày trong Hình 8.

Các thực nghiệm được mô tả theo Bảng 2 đã cho thấy rằng sự kết hợp giữa phương pháp k-Means và cấu trúc R^S-Tree trong việc tìm kiếm hình ảnh đã đem lại sự cải thiện về độ chính xác. Cụ thể, độ chính xác trung bình của quá trình tìm kiếm ảnh theo nội dung trên cấu trúc KMRST của các bộ dữ liệu ảnh COREL, OxfordFlowers-17 và MS-COCO lần lượt là 79,54% (tăng 4,26%), 82,38% (tăng 4,69%) và 70,96% (tăng 6,90%), đã tăng hiệu suất so với việc tìm kiếm trên cấu trúc R^S-Tree. Độ chính xác tăng lên là do cấu trúc KMRST đã áp dụng thuật toán k-Mean tại tầng nút lá của cấu trúc R^S-Tree để thực hiện việc phân cụm lại các nút lá. Quá trình này giúp cho các phần tử được phân bố một cách chính xác hơn vào các cụm tương ứng của chúng. Đồng thời, việc kết hợp cấu trúc KMRST với Ontology cũng đã đem lại kết quả ấn tượng trong việc tìm kiếm hình ảnh theo tiếp cận ngữ nghĩa. Sự kết hợp này đã giúp tăng độ chính xác đáng kể so với việc chỉ tìm kiếm theo nội dung. Cụ thể, độ chính xác trung bình của việc tìm kiếm hình ảnh theo ngữ nghĩa trên mô hình đề xuất lần lượt là 95,71% (tăng 20,33%), 94,63% (tăng 14,87%) và 81,57% (tăng 14,95%), hiệu suất tăng lên so với việc tìm kiếm theo nội dung trên cấu trúc KMRST. Từ những kết quả thực nghiệm này, chúng ta có thể kết luận rằng việc kết hợp cấu trúc KMRST và phương pháp tìm kiếm dựa trên ngữ nghĩa đã mang lại hiệu suất tốt trong việc tìm kiếm hình ảnh tương tự theo cách tiếp cận mới.

Bảng 2. Hiệu suất tìm kiếm của các phương pháp trên các tập dữ liệu ảnh

Phương pháp	Precision (MAP %)	Recall (MAP %)	F-measure (MAP %)	Query time (mili giây)	Bộ ảnh
RST_CBIR	76,29	68,75	72,32	14,39	COREL
KMRST_CBIR	79,54	70,66	74,84	17,33	COREL
KMRSTOnIR	95,71	89,75	92,63	26,28	COREL
RST_CBIR	78,69	71,69	75,03	11,45	OxfordFlowers-17
KMRST_CBIR	82,38	73,06	77,44	15,64	OxfordFlowers-17
KMRSTOnIR	94,63	88,76	91,60	24,95	OxfordFlowers-17
RST_CBIR	66,38	58,75	62,33	69,47	MS-COCO
KMRST_CBIR	70,96	62,33	66,37	70,97	MS-COCO
KMRSTOnIR	81,57	76,54	78,97	80,55	MS-COCO



Hình 8. Biểu đồ so sánh hiệu suất các phương pháp tìm kiếm ảnh trên các bộ dữ liệu

Bảng 3. So sánh độ chính xác tìm kiếm ảnh trên các bộ dữ liệu ảnh

Phương pháp	MAP(%)	Bộ dữ liệu
Shape and Texture Features, 2019 [1]	72,86	COREL
Multi-feature and SVM, 2018 [7]	76,57	COREL
Multi-feature with neural network, 2020 [17]	79,41	COREL
KMRSTOnIR	95,71	COREL
CNN+SURF+BIC, 2019 [8]	85,39	OxfordFlowers17
Bag-of-Words, 2018 [2]	77,10	OxfordFlowers17
Dictionary Learning, 2013 [6]	73,43	OxfordFlowers17
KMRSTOnIR	94,63	OxfordFlowers17
CAMP, 2019 [21]	68,90	MS-COCO
Hamming Ranking using AlexNet, 2021 [25]	76,40	MS-COCO
Resnet, 2020 [22]	81,10	MS-COCO
KMRSTOnIR	81,57	MS-COCO

Bảng 3 thể hiện kết quả so sánh giữa mô hình đề xuất với các nghiên cứu gần đây trên các bộ dữ liệu COREL, OxfordFlowers-17 và MS-COCO. Đối với tập ảnh COREL mô hình đề xuất có độ chính xác là 95,71%. Đối với tập dữ liệu OxfordFlowers-17, độ chính xác là 94,63%. Đối với tập ảnh MS-COCO, có độ chính xác là 81,57%. Trên cơ sở số liệu của các bảng so sánh này cho thấy phương pháp đề xuất cao hơn các công trình nghiên cứu liên quan. Điều này cho thấy rằng phương pháp đề xuất là khả thi và hiệu quả để áp dụng vào các hệ tìm kiếm ảnh.

VI. KẾT LUẬN

Bài báo này đã tiếp cận sử dụng mô hình tìm kiếm ảnh dựa trên sự kết hợp kỹ thuật k-Means và cấu trúc R^S -Tree; đồng thời sử dụng mạng nơ-ron tích chập kết hợp với cấu trúc túi từ thị giác để phân bố hình ảnh vào khung Ontology. Mô hình truy vấn ảnh được đề xuất là sự kết hợp giữa truy vấn theo nội dung và theo tiếp cận ngữ nghĩa, trong mô hình này được thiết kế gồm 2 phần: (1) kết hợp kỹ thuật k-Means với cấu trúc dữ liệu R^S -Tree, đồng thời học huấn luyện mạng R-CNN và làm giàu Ontology dựa trên việc phân bố các hình ảnh trong túi từ thị giác; (2) thực hiện quá trình tìm kiếm ảnh dựa trên đặc trưng cấp thấp sử dụng cấu trúc KMRST và tìm kiếm ảnh theo tiếp cận ngữ nghĩa trên Ontology. Việc kết hợp tìm kiếm theo nội dung kết hợp với ngữ nghĩa đã nâng cao hiệu quả cho bài toán tìm kiếm ảnh, đồng thời đáp ứng được cho các tập dữ liệu tăng trưởng. Mô hình đã được thử nghiệm trên các bộ ảnh COREL, OxfordFlowers17 và MS-COCO, đạt được độ chính xác lần lượt là 95,71%, 94,63% và 81,57%. So sánh với các nghiên cứu khác, mô hình đề xuất đã cho thấy tính hiệu quả trong việc giải quyết bài toán tìm kiếm ảnh tương tự. Mô hình tìm kiếm ảnh đề xuất dựa trên nội dung và đồng thời tập trung vào việc giảm khoảng cách ngữ nghĩa thông qua việc kết nối và truy vấn trên Ontology. Tuy nhiên, việc trích xuất các chú giải và ngữ nghĩa đối tượng vẫn chưa được thực hiện do yêu cầu phải làm giàu dữ liệu từ ngữ cho Ontology. Do đó, hướng phát triển tiếp theo sẽ xây dựng hệ thống truy vấn ảnh dựa trên ngữ nghĩa, cho phép trích xuất trực tiếp các mô tả hình ảnh thông qua đồ thị tri thức. Mục tiêu của hướng phát triển này là làm giảm khoảng cách ngữ nghĩa, đồng thời tiến gần hơn đến ngôn ngữ tự nhiên mà con người sử dụng.

TÀI LIỆU THAM KHẢO

- [1] Abdulkadhem, A.A., and Tawfiq A. Al-Assadi, "Proposed a Content-Based Image Retrieval System Based on the Shape and Texture Features," *Int. J. Innovat. Technol. Explor. Eng.*, 2019. 8 (2189).
- [2] Ahmed, K.T., S. Ummesafi, and A. Iqbal, "Content based image retrieval using image features information fusion," *Information Fusion*, 2019. 51: pp. 76-99.
- [3] Al-Mohamade, A., O. Bchir, and M.M. Ben Ismail, "Multiple query content-based image retrieval using relevance feature weight learning," *Journal of imaging*, 2020. 6 (1): pp. 2.
- [4] Asim, M.N., et al., "The use of Ontology in retrieval: a study on textual, multilingual, and multimedia retrieval," *IEEE Access*, 2019. 7: pp. 21662-21686.
- [5] Bansal, M., et al., "An efficient technique for object recognition using Shi-Tomasi corner detection algorithm," *Soft Computing*, 2021. 25: pp. 4423-4432.
- [6] Gao, S., I.W.-H. Tsang, and Y. Ma, "Learning category-specific dictionary and shared dictionary for fine-grained image categorization," *IEEE Transactions on Image Processing*, 2013. 23 (2): pp. 623-634.
- [7] Gonçalves F. M. F., G.I.R., Pedronette D. C. G., "Semantic guided interactive image retrieval for plant identification," *Expert Systems with Applications*, 2018(91): pp. 12-26.
- [8] Gonçalves, F.M.F., I.R. Guilherme, and D.C.G. Pedronette, "Semantic guided interactive image retrieval for plant identification," *Expert Systems with Applications*, 2018. 91: pp. 12-26.

- [9] Gupta, D., et al., "Medical Image Retrieval via Nearest Neighbor Search on Pre-trained Image Features," *arXiv preprint arXiv:2210.02401*, 2022.
- [10] Joseph, A., et al., "Content-based image retrieval using hybrid k-means moth flame optimization algorithm," *Arabian Journal of Geosciences*, 2021. 14 (8): pp. 687.
- [11] Khan, U.A., A. Javed, and R. Ashraf, "An effective hybrid framework for content based image retrieval (CBIR)," *Multimedia Tools and Applications*, 2021. 80: pp. 26911-26937.
- [12] Le, T.M., "An improvement of r-tree for content-based image retrieval," in *Annales Universitatis Scientiarum Budapestinensis de Rolando Eotvos Nominatae. Sectio Computatorica*, 2022.
- [13] Liu, Y., et al., "Integrating object Ontology and region semantic template for crime scene investigation image retrieval," in *2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. 2017. IEEE.
- [14] Manzoor, U., et al., "Semantic image retrieval: An Ontology based approach," *International Journal of Advanced Research in Artificial Intelligence*, 2015. 4 (4): pp. 1-8.
- [15] Maur, H.K., P. Faridkot, and P. Jain, "Content based image retrieval system using K-means clustering algorithm and SVM classifier technique," 2019.
- [16] Murthy, V., et al., "Application of hierarchical and K-means techniques in Content based image retrieval," *International Journal of Engineering Science and Technology*, 2010. 2 (5): pp. 749-755.
- [17] Raja R., K.S., Mahmood M. R., "Color object detection based image retrieval using ROI segmentation with multi-feature method," *Wireless Personal Communications*, 2020: pp. 169-192.
- [18] Rejito, J., et al., "Image indexing using color histogram and k-means clustering for optimization CBIR in image database," in *Journal of Physics: Conference Series*. 2017. IOP Publishing.
- [19] Sivakumar, M., N.S. Kumar, and N. Karthikeyan, "Content-based image retrieval techniques: a survey," in *Journal of Physics: Conference Series*. 2021. IOP Publishing.
- [20] Sulaiman, M.S., S. Nordin, and N. Jamil, *An object properties filter for multi-modality Ontology semantic image retrieval*. Journal of Information and Communication Technology, 2017. 16 (1): pp. 1-19.
- [21] Wang, A., Wang, Y., & Chen, Y., *Hyperspectral image classification based on convolutional neural network and random forest*. Remote sensing letters, 2019. 10 (11): pp. 1086-1094.
- [22] Wen, S., Liu, W., Yang, Y., Zhou, P., Guo, Z., Yan, Z., ... & Huang, T., "Multilabel image classification via feature/label co-projection," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 51, 2020.
- [23] White, D.A. and R. Jain., "Similarity indexing with the SS-tree," in *Proceedings of the Twelfth International Conference on Data Engineering*. 1996. IEEE.
- [24] Younus, Z.S., et al., "Content-based image retrieval using PSO and k-Means clustering algorithm," *Arabian Journal of Geosciences*, 2015. 8: pp. 6211-6224.
- [25] Zhang, Z., Peng, A., & Li, H., "Instance-weighted central similarity for multi-label image retrieval," *arXiv preprint arXiv:2108*, 2018.

COMBINING K-MEANS WITH R^S-TREE FOR SEMANTIC-BASED IMAGE RETRIEVAL

Le Thi Vinh Thanh, Le Manh Thanh, Van The Thanh, Nguyen Thi Uyen Nhi

ABSTRACT: In this paper, a structure combining k-Means with R^S-Tree for image retrieval is proposed, named KMRST. The construction process of this structure consists of two phases: (1) R^S-Tree structure construction phase, where the k-Means algorithm is utilized to split nodes into k-clusters; (2) R^S-Tree structure completion phase, where the k-Means algorithm is employed at leaf nodes to re-cluster the elements. Additionally, an Ontology is built based on visual bag-of-words to describe the semantics and relationships among objects in the images. For a given input image, the retrieval process is conducted on the constructed KMRST structure to find a set of similar images. Subsequently, the input image is classified using the k-NN method based on the set of similar images, generating a visual vocabulary. Finally, a SPARQL query is generated from this vocabulary and performs semantic-based image retrieval on the Ontology. Building on the established theory, a semantic image retrieval model is proposed. Experiments are conducted on image datasets including COREL, OxfordFlowers17, and MS-COCO. The experimental results are evaluated and compared with other works to demonstrate the effectiveness of the proposed method.

Keywords: SBIR, Image retrieval, R^S-Tree, k-Means, Ontology.